双字词识别中首、尾词素位置概率对位置编码灵活程度的影响

李馨[1]，张美[1]，顾俊娟[2]，王永胜[1]，梁菲菲[1*]

（[1]天津师范大学心理学部，天津，300387；[2]华北理工大学心理与精神卫生学院，唐山，063210）

摘　要 本研究通过两个平行实验，探讨首、尾词素位置概率如何影响双字词识别中位置编码的灵活程度及其时间进程。采用掩蔽启动词汇判断任务，操纵启动条件（原词启动、转换启动、替换启动）和启动时间（80ms、150ms、300ms）。实验 1 操纵首词素位置概率高低，保证尾词素相同；实验 2 操纵尾词素位置概率高低，保证首词素相同。以 144 名大学生为研究对象，结果发现：（1）词汇位置编码的灵活性体现在词汇识别全程。（2）首、尾词素位置概率均作用于词汇识别，但尾词素位置概率的作用时程更长。（3）首词素而不是尾词素的位置概率调节词汇位置编码的灵活程度，当启动时间为 150ms，首词素经常用在词首时词汇的位置编码更灵活。由此推断，首词素的位置概率可能与词汇识别的位置编码灵活程度相关；尾词素的位置概率则与位置编码的灵活程度无关。

关键词 词素转置效应，词素位置概率，词汇位置编码，双字词识别

分类号 B842

1 引言

词汇识别是阅读理解的基础。成功的词汇识别涉及两类信息加工：第一，身份信息加工（Identity information），在“what”层面回答词汇由哪些字符组成；第二，位置信息加工（Position information），在“where”层面回答字符在词汇中的相对顺序（滑慧敏 等, 2017; 徐迩嘉, 隋雪, 2018; 张妍萃 等, 2021）。中文词汇长度相对较短，约 60%的词汇是双字词（Hansell, 2008），这使得词素的位置信息在词汇识别中显得尤为重要。特别是转置词（构成词汇的身份信息相同，但位置信息和语义不同，如“领带”和“带领”）的存在，使得中文双字词的位置编码机制显得更为复杂。已有研究发现，中文词汇识别中的位置编码并非完全严格，而是具有一定的灵活性（Gu et al., 2015）。然而，目前尚不清楚中文双字词识别中的位置编码灵活程度受哪些因素影响，以及在什么阶段受到影响？本研究通过两项平行实验，从加工进程视角考察首、尾词素位置概率高低如何影响双字词位置编码的灵活程度，从而深化中文词汇识别中位置信息加工的认识。

收稿日期: 2024-06-21

* 教育部人文社会科学规划一般项目(21YJA190004)。

通信作者: 梁菲菲, E-mail: feifeiliang_329@126.com

在词汇识别过程中，字母/词素位置编码的灵活性已在多项研究中证实。例如，在掩蔽启动范式和边界范式的研究中，转置字符非词（单词中的两个字母位置互换，如 jugde）相较于替换字符非词（单词中的两个字母替换为其它字符，如 jupte），更能有效促进基词（如 judge）的激活。这一现象表明，虽然字母位置对词汇识别至关重要，但是在某些条件下，能够在一定程度上容忍字母位置的变化，这一现象在许多拼音文字中（如英语、法语、西班牙语、泰语等）得到证实，被称为字母转置效应（transposed-letter effect，简称 TL 效应，见 Baciero et al., 2022; Blythe et al., 2014; Johnson et al., 2007; Perea & Fraga, 2006; Perea et al., 2012）。由于中文词汇由词素构成，由此称之为词素转置效应。大多数拼音文字语言属于线性文字系统（如英语、法语等），词汇由单个或多个字母构成，这些字母具有形、音信息而缺乏语义信息。此外，由于拼音文字语言的正字法深度较浅，形-音对应较为系统，其字母转置效应主要与字母形态有关。空间编码模型（Spatial Coding Model, 见 Davis, 2010）和重叠模型（The Overlap Model, 见 Gomez et al., 2008）从视觉加工层面为字母转置效应提供了理论支持。这些模型指出，受视觉系统限制，词汇加工早期的字母位置分布较广，具有不确定性，相比于替换字母非词，转置字母非词（如 jugde）与原词的字母位置分布（如 judge）更相似、空间激活梯度模式更相近，因此更容易被加工为原词，由此表现出字母位置加工的灵活性。

相比之下，中文词汇识别的词素转置效应机制可能更为复杂。首先，中文词汇在长度上短于英文词汇（多数词汇由两个汉字构成），且每个汉字不仅携带形、音、义信息，还承载着关于词汇的独特语义。就双字词举例，当两个汉字的位置发生调换时，相当于构成整词的所有词素位置均发生改变，转置非词（如“母字”）与原词（如“字母”）在空间激活模式上的差异较大；相比之下，由于拼音文字的词汇较长（包含多个字母），且每个字母只具有形、音信息，调换其中两个字母的位置后，转置非词与原词在空间激活模式上的差异相对较小。由此推断，中文词汇识别中的位置编码在灵活程度上可能有别于拼音文字。其次，由于中文词汇基于词素构成，每个词素都是表达语义的最小单位。在这种情况下，转置非词的加工在涉及词素位置识别错误时，可能会受到词素本身特征的影响，如词素的视觉复杂性（梁菲菲, 刘瑛, 贺斐, 冯琳琳, 王峥, 白学军, 2024）、使用频率（宋悉妮 等, 2022）以及本身所携带的位置概率信息（Liang et al., 2023）等。由于中文词汇识别中的词素转置效应较为复杂，目前关于位置编码灵活性还没有形成一个合理解释。

我们尝试运用中文阅读模型（Chinese Reading Model，简称 CRM，见 Li & Pollatsek, 2020）的基本观点，对中文词汇位置灵活编码的机制做出假设。该模型主张，词汇加工模块包括三

个层级：视觉层、汉字层和词汇层。三个加工层级中的编码方式具有位置特异性。具体来说，在汉字加工层级，知觉广度内的所有汉字被平行激活，由该汉字组成的所有词汇也会被不同程度的激活。例如，当“大”被激活时，与“大”相关的词汇也会被激活，有时“大”位于词首（如“大人”、“大学”等），有时“大”位于词尾（如“博大”、“壮大”等）。也就是说，在汉字加工层级，“大”字身份信息和位置信息逐步得到加工，由“大”组成的所有词汇均被不同程度的激活。由于所激活的词汇中“大”位于词内的不同位置，此时对“大”的位置编码是相对灵活的。随着词汇身份信息的激活增强，汉字的位置信息逐渐被固定，位置编码灵活程度开始降低。如果该假设正确，我们推测，中文双字词识别中的位置编码灵活程度可能受词素的位置概率高低影响。

词素位置概率信息是指汉字出现在多字词特定位置的概率（如词首、词中、词尾，见连坤予 等, 2021; 梁菲菲, 冯琳琳, 刘瑛, 李馨, 白学军, 2024; Liang et al., 2023）。例如，“彼”字共组成 12 个词（如“彼岸”、“彼此”等），“彼”在其中均位于词首，则“彼”位于词首的概率为 100%，位于词尾的概率为 0。词素位置概率信息是中文的一种语言学信息，可能作用于词汇识别的汉字加工层级（见曹海波 等, 2023; 梁菲菲, 冯琳琳, 刘瑛, 李馨, 白学军, 2024; Liang et al., 2015, 2017, 2023）。Liang 等人（2015）提出，当知觉广度范围内汉字构成的词汇被激活时，该汉字在多字词中的位置概率信息同时被激活，当词素在词内的实际位置与其常用位置一致时，词素处于该位置的激活程度就越高，则容易被识别；如果词素实际位置与其位置概率信息不符，会造成认知冲突，读者需要使用额外认知资源解决这一冲突，导致词汇识别速度较慢。既然词素位置概率信息加工和词汇的位置加工相关，那么词素位置概率的高低如何影响词汇位置编码的灵活性？这是本研究的主要问题。

此外，以往研究发现，首、尾词素的特征在中文词汇识别中的作用及加工方式并不完全相同。大多数研究支持首词素识别优势的观点，证据如下：第一，首、尾词素的字频均影响词汇识别，但只有首词素为高频字时，尾词素的字频才发挥作用（Yan et al., 2006）；第二，首词素的视觉复杂性同时影响词汇识别和眼跳定位，而尾词素的视觉复杂性只影响词汇识别（Ma & Li, 2015）；第三，首、尾词素的位置概率信息均作用于阅读伴随词汇学习的词切分，但首词素的作用时程更长，更稳定（梁菲菲, 冯琳琳, 刘瑛, 李馨, 白学军, 2024）。中文阅读的首词素识别优势可能是因为阅读过程中的视觉词汇识别方向为从左到右，读者对首词素的加工先于尾词素，导致首词素在词汇识别中起关键作用。由此，我们推测，首词素由于词内位置的特殊性，可能对词汇识别中的位置灵活编码的影响更大；而尾词素的位置相对不重要，可能对词汇识别中的位置灵活编码的影响相对较小。本研究将分别考察首、尾词素的位置概

率对词汇位置编码灵活程度的调节，以检验该假设是否正确。

关于词汇位置灵活编码发生在词汇识别的哪一阶段，还存在争议。目前有两种观点：第一，词汇位置灵活编码发生在词汇识别中期。彭聃龄等人（1999）采用语义启动范式，以中文逆序词（如“领带”、“带领”）为实验材料，操纵启动类型（原词启动、逆序词启动、无关启动）和 SOA（57ms，157ms，314ms），目标词为原词的语义相关词。结果发现，当 SOA 为 157ms 时，原词和逆序词的语义启动效应无显著差异，均促进目标词加工，表明此时词汇位置编码是灵活的；而在加工早期（SOA 为 57ms）和晚期（SOA 为 314ms），逆序词的启动效应显著小于原词启动效应，表明此时位置编码不灵活。上述研究结果表明词汇位置编码经历了“不灵活-灵活-不灵活”的过程，在词汇加工中期，词汇位置编码的灵活程度最高。第二，词汇位置灵活编码发生在词汇识别全程。徐迩嘉和隋雪（2018）采用掩蔽启动范式，发现在不同启动时间下（80ms、150ms、300ms），均存在词素转置效应，表明词汇位置灵活编码稳定地存在于词汇识别过程中，但启动时间为 150ms 时词素转置效应量最大，即在词汇识别中期位置编码最灵活。基于词素位置灵活编码的发生阶段还无统一定论，本研究有必要通过操纵不同的启动时间（80ms、150ms、300ms），探究首、尾词素位置概率影响词汇位置编码灵活程度的时间进程。

本研究包括两项平行实验：实验一操纵首词素位置概率高低，实验二操纵尾词素位置概率高低，设置三种启动条件（原词启动、转换启动、替换启动）和三个启动时间（80ms，150ms，300ms），分别考察首、尾词素位置概率信息高低如何影响词汇位置编码的灵活程度及其时间进程。研究假设为：第一，基于启动 150ms 时的词汇位置编码灵活性最强（彭聃龄 等, 1999; 徐迩嘉, 隋雪, 2018），我们预期，词素位置概率信息最有可能在启动时间为 150ms 时影响词汇位置编码的灵活程度（用词素转置效应量来衡量）。第二，基于首词素的位置优势，以及在双字词识别中的加工优势（Ma & Li, 2015; Yan et al., 2006），我们预期：首词素的位置概率对词汇位置编码灵活程度的影响大于尾词素。

2 实验 1　首词素位置概率对词汇位置编码灵活程度的影响及时间进程

2.1 实验目的

采用掩蔽启动词汇判断任务，操纵首词素的位置概率和启动类型，考察首词素位置概率如何影响词汇位置编码的灵活程度。

2.2 方法

2.2.1 被试

实验 1a（启动 80ms），实验 1b（启动 150ms）和实验 1c（启动 300ms）各包括 48 名天津师范大学在校生，共 144 名被试，平均年龄为 20.21 岁（*SD* = 1.74）。所有被试的母语均为汉语，视力或矫正视力正常，均不知晓实验目的，实验结束后给予一定报酬。

采用 G*power 估算样本量。三个子实验均为 2×3 被试内实验设计，在显著性水平α = 0.01 且中等效应量水平（f = 0.25）时，预测样本量至少为 36，可使统计检验力达到 0.95。三个子实验的被试数量均为 48 人，超过最小样本量。

2.2.2 实验设计

三个实验（启动时间分别为 80ms, 150ms 和 300ms）均为 2（首词素位置概率：高、低）×3（启动类型：原词启动、转换启动、替换启动）被试内实验设计。因变量为正确率和反应时。

2.2.3 实验材料

参照前人研究（Liang et al., 2015, 2017），将词素位置概率高于 70%定义为高概率条件，低于 30%定义为低概率条件。从 SUBTLEX-CH 语料库（Cai & Brysbaert, 2010）中选取 126 对低频且不可逆的双字词（平均词频：1.11 次/百万；词频范围：0.03 次/百万~5.72 次/百万）作为目标词。操纵首词素位置概率的高低：高概率条件下，首词素用在双字词词首的平均概率为 85.35%（70.49%~100%）；低概率条件下，首词素用在双字词词首的平均概率为 22.78%（5.16%~29.73%）。保证每对目标词的尾词素相同，且尾词素位置概率处于中等水平（40%~60%）。匹配两类目标词的词频、首词素字频、首词素笔画数、整词笔画数。配对样本 t 检验结果显示，两类目标词的词频、首词素字频、首词素笔画数、整词笔画数均无显著差异，*t*s < 1，*p*s > 0.05。目标词材料的基本特征统计见表 1。

表 1 目标词材料基本特征统计

属性	高首位置概率目标词	低首位置概率目标词	转换启动词	替换启动词
整词词频	1.15(1.26)	1.07(1.26)	—	—
整词笔画数	17.06(4.64)	17.03(3.68)	17.04(4.18)	16.91(4.11)
首字字频	545.24(2066.44)	561.56(1187.20)	519.27(799.86)	497.64(824.75)
首字笔画数	8.71(3.35)	8.68(2.89)	8.35(2.85)	8.23(2.88)
尾字字频	519.27(801.46)	519.27(801.46)	553.40(1681.83)	504.08(1257.42)
尾字笔画数	8.35(2.86)	8.35(2.86)	8.69(3.12)	8.68(2.94)

注：括号内为标准差，以下同。

每个目标词有三种启动条件：（1）原词条件，启动词与目标词相同，如“值班—值班”；（2）转换条件，启动词为目标词的两个词素换位形成的假词，如“班值—值班”；（3）替换条件，启动词为不同于目标词词素的两个汉字组成的假词，如“街旅—值班”。转换条件和替换条件下启动词的两个汉字在结构上匹配。同时，转换启动词的首字字频和笔画数与替换启动词的首字无显著差异，$ts < 1.11$，$ps > 0.05$。转换启动词的尾字字频和笔画数与替换启动词的尾字无显著差异，$ts < 1$，$ps > 0.05$。

为平衡词汇判断任务中“是”、“否”反应数量，加入 252 个假词作为填充。填充词包括三种启动条件，与目标词的操纵相同（例如，“资流—资流”，“流资—资流”，“烦背—资流”）。目标词的首词素字频和笔画数（字频：$M = 553.40$ 次/百万，$SD = 1681.83$ 次/百万；笔画数：$M = 8.69$，$SD = 3.12$）与填充词无显著差异（字频：$M = 542.96$ 次/百万，$SD = 1827.41$ 次/百万；笔画数：$M = 8.68$，$SD = 2.94$；$ts < 1$，$ps > 0.05$）。目标词的尾词素字频和笔画数（字频：$M = 519.27$ 次/百万，$SD = 799.86$ 次/百万；笔画数：$M = 8.35$，$SD = 2.85$）与填充词无显著差异（字频：$M = 521.41$ 次/百万，$SD = 815.30$ 次/百万；笔画数：$M = 8.57$，$SD = 2.73$；$ts < 1.01$，$ps > 0.05$）。填充词材料的基本特征统计见表 2。

表 2 填充词材料基本特征统计

属性	填充词	目标词	转换启动词	替换启动词
首字字频	542.96(1827.41)	553.40(1681.31)	521.41(815.30)	523.68(1066.69)
首字笔画数	8.68(2.94)	8.69(3.12)	8.57(2.73)	8.67(2.82)
尾字字频	521.41(815.30)	519.27(799.86)	542.96(1827.41)	524.94(1068.47)
尾字笔画数	8.57(2.73)	8.35(2.85)	8.68(2.94)	8.60(2.76)
总笔画数	17.29(4.00)	17.04(4.18)	17.29(4.00)	17.27(3.84)

为检验假词操纵的有效性，对目标词的转换启动词和替换启动词、填充词以及填充词的转换启动词和替换启动词进行评定，共 1250 个假词。由于假词数量较多，分为 4 个版本，每个版本有 10 人填写，共 40 名大学生。采用迫选法，要求被试从两个选项（一个真词和一个假词）中选出假词，5 种条件的作答正确率均大于 95%，表明假词操纵是有效的。

为确保被试在不同启动条件下不会看到同一目标词，将实验材料进行拉丁方分组，分为 3 个 block，每个 block 中包含 252 个目标词和 252 个填充词，每个启动条件下各 84 个目标词。每名被试完成一个 block 的实验。

2.2.4 实验仪器

采用联想昭阳 k4e-itl 笔记本电脑，屏幕为 14 英寸，分辨率为 1024×728 像素。实验材料以白底黑字呈现。启动词和目标词均为 40 号宋体加粗，眼睛与屏幕之间的距离为 60cm，

每个汉字所占的视角约为 1.9°。

2.2.5 实验程序

实验使用 E-prime3.0 编程。被试单独施测。实验开始时，屏幕中央呈现“+”注视点，500ms 后注视点消失，呈现前掩蔽刺激“**”500ms，接着呈现启动词 80ms，随后呈现后掩蔽刺激“**”40ms，之后出现目标词，要求被试又快又准地进行真假词判断，真词按“F”键，假词按“J”键。被试按键反应后，目标词消失，随后空屏 1000ms 进入下一试次。实验材料随机呈现。正式实验前有 8 个练习试次。如果被试能够正确理解实验任务，可选择开始正式实验，否则再次返回练习阶段。在正式实验阶段，每个被试完成 504 个试次。实验需 20 分钟左右。实验流程见图 1。

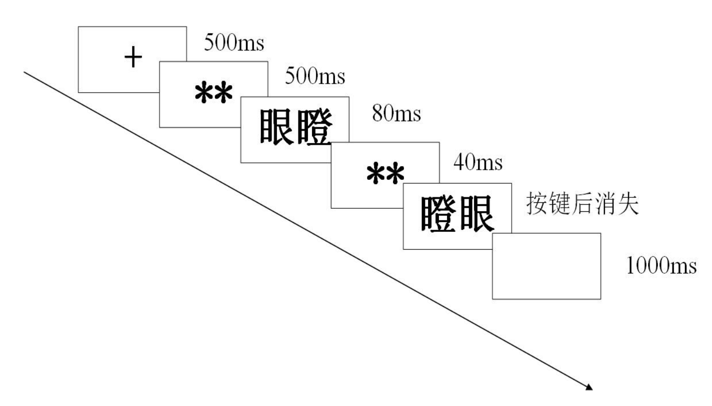


图 1 实验 1a（启动 80ms）流程图

注：以转换启动条件为例，启动词为“眼瞪”，目标词为“瞪眼”。

2.3 结果

参照 Gu 等人（2015）研究，删除反应时小于 100ms、大于 1500ms，以及每个被试在每种实验条件下大于或小于 3 个标准差的数据，3 个实验中删除数据分别占总数据的 6.14%、4.73%、4.80%。填充材料不参与分析。

使用 R 统计软件（R Development Core Team, 2016）和 lme4 工具包（Bates et al., 2023），采用广义线性混合模型（Generalized Linear Mixed-effects Models, GLMMs）和线性混合模型（Linear Mixed-effects Models, LMMs）分别对正确率和反应时进行分析。将反应时 log 转换，首词素位置概率、启动类型、二者的交互作用作为固定效应，被试、项目作为随机效应纳入模型。将启动类型中的转换启动条件作为基线，事先设置两个比较：第一，原词启动与转换启动比较，反映词汇位置编码的严格程度；第二，替换启动与转换启动比较，得到词素转置效应，反映位置编码的灵活性。由于本研究主要考察词汇位置编码的灵活性，即词素转置效应，因此，在结果部分仅报告第二种比较的结果。采用最大随机效应结构模型，若无法拟合，则采用逐渐递减原则，直至模型拟合成功。

2.3.1 实验 1a（启动时间为 80ms）的正确率和反应时分析

各个实验条件下的正确率和反应时结果见表 3 所示，模型统计结果汇总见表 4 所示。

正确率分析。首词素位置概率主效应不显著（b = -0.29, SE = 0.16, z = -1.84, p = 0.07），转换启动和替换启动的正确率差异不显著（b = -0.20, SE = 0.19, z = -1.08, p = 0.28）；交互作用不显著（b = 0.34, SE = 0.38, z = 0.89, p = 0.38）。

反应时分析。首词素位置概率主效应不显著（b = 0.01, SE = 0.01, t = 1.01, p = 0.32），表明当启动时间为 80ms 时，首词素的位置概率信息未得到加工。替换启动的反应时显著长于转换启动（b = 0.06, SE = 0.01, t = 6.21, p ＜ 0.001），表明此时位置编码是灵活的。交互作用不显著（b = -0.01, SE = 0.01, t = -0.40, p = 0.69）。

上述结果表明，启动时间为 80ms 时，词汇位置编码具有灵活性。但由于首词素位置概率信息未得到加工，因此不会调节词汇位置编码的灵活程度。

表 3 启动 80ms 时不同实验条件下的正确率和平均反应时（ms）

	正确率		反应时(ms)	
启动类型	高首词素位置概率	低首词素位置概率	高首词素位置概率	低首词素位置概率
原词启动	0.94(0.05)	0.92(0.05)	625(98)	622(94)
转换启动	0.94(0.05)	0.90(0.06)	708(99)	719(103)
替换启动	0.92(0.07)	0.90(0.08)	749(95)	755(98)

表 4 启动 80ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	*b*	*SE*	*t/z*	95%CI
正确率	截距	3.00	0.12	25.94	[2.78, 3.23]
	首词素位置概率	-0.29	0.16	-1.84	[-0.60, 0.02]
	启动类型（UN vs TC）	-0.20	0.19	-1.08	[-0.58, 0.17]
	首词素位置概率×启动类型（UN vs TC）	0.34	0.38	0.89	[-0.41, 1.08]
反应时	截距	6.51	0.02	341.63***	[6.47, 6.54]
	首词素位置概率	0.01	0.01	1.01	[-0.01, 0.02]
	启动类型（UN vs TC）	**0.06**	**0.01**	**6.21*****	**[0.04, 0.08]**
	首词素位置概率×启动类型（UN vs TC）	-0.01	0.01	-0.40	[-0.04, 0.03]

注：*p < 0.05，**p < 0.01，***p < 0.001，CI = Confidence Interval（置信区间）。TC =转换启动；UN = 替换启动。以下同。

2.3.2 实验 1b（启动时间为 150ms）的正确率和反应时分析

各个实验条件下的正确率和反应时结果见表 5 所示，模型统计结果汇总见表 6 所示。

正确率分析。首词素位置概率主效应边缘显著（b = -0.22, SE = 0.12, z = -1.89, p =

0.06），高首词素位置概率条件下的正确率有高于低首词素位置概率条件的趋势；转换启动条件的正确率显著高于替换启动（b = -0.60, SE = 0.14, z = -4.22, p ＜ 0.001）；交互作用不显著（b = -0.10, SE =0.28, z = -0.34, p = 0.69）。

反应时分析。首词素位置概率主效应显著（b = 0.02, SE = 0.01, t = 2.89, p = 0.004），低首词素位置概率词的反应时显著长于高首词素位置概率词；替换启动的反应时显著长于转换启动（b = -0.01, SE = 0.01, t = 11.61, p ＜ 0.001）；首词素位置概率与启动类型（替换启动 vs. 转换启动）的交互作用显著（b = -0.03, SE = 0.01, t = -2.28, p = 0.02）。简单效应分析发现，两类目标词均表现出显著的词素转置效应（高首词素位置概率：b = 0.11, SE = 0.03, t = 3.70, p ＜ 0.001; 低首词素位置概率：b = 0.08, SE = 0.03, t = 2.71, p = 0.007）。相比之下，高首词素位置概率词的词素转置效应（71ms）显著大于低首词素位置概率词（50ms），表明，当目标词的首词素经常用在词首时，词汇位置编码灵活程度更大。

上述结果表明，启动 150ms 时，首词素位置概率信息得到加工，且调节词汇位置编码的灵活程度，当双字词的首字经常用在词首时，位置编码灵活程度更高。

表 5 启动 150ms 时不同实验条件下的正确率和平均反应时（ms）

	正确率		反应时(ms)	
启动类型	高首词素位置概率	低首词素位置概率	高首词素位置概率	低首词素位置概率
原词启动	0.94(0.05)	0.93(0.06)	527(89)	530(89)
转换启动	0.94(0.05)	0.93(0.04)	644(96)	669(95)
替换启动	0.91(0.06)	0.89(0.08)	715(82)	719(85)

表 6 启动 150ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	*b*	*SE*	*t/z*	95%CI
正确率	截距	2.85	0.10	28.87	[2.67, 3.04]
	首词素位置概率	-0.22	0.12	-1.89	[-0.45, 0.01]
	启动类型（UN vs TC）	**-0.60**	**0.14**	**-4.22*******	**[-0.88, -0.32]**
	首词素位置概率×启动类型（UN vs TC）	-0.10	0.28	-0.34	[-0.65, 0.46]
反应时	截距	6.41	0.02	352.30***	[6.37, 6.44]
	首词素位置概率	**0.02**	**0.01**	**2.89****	**[0.01, 0.03]**
	启动类型（UN vs TC）	**0.10**	**0.01**	**11.61*******	**[0.08, 0.12]**
	首词素位置概率×启动类型（UN vs TC）	**-0.03**	**0.01**	**-2.28***	**[-0.06, -0.004]**

2.3.3 实验 1c（启动时间为 300ms）的正确率和反应时分析

各个实验条件下的正确率和反应时结果见表 7 所示，模型统计结果汇总见表 8 所示。

正确率分析。首词素位置概率主效应不显著（b = -0.20, SE = 0.14, z = -1.44, p = 0.15）；

转换启动的正确率显著高于替换启动（$b = -0.48, SE = 0.17, z = -2.88, p = 0.004$）；交互作用不显著（$b = 0.23, SE = 0.33, z = 0.68, p = 0.50$）。

反应时分析。首词素位置概率主效应不显著（$b = 0.01, SE = 0.01, t = 0.80, p = 0.43$）；替换启动的反应时显著长于转换启动（$b = 0.13, SE = 0.01, t = 11.06, p < 0.001$）；交互作用不显著（$b = -0.01, SE = 0.02, t = -0.71, p = 0.48$）。

上述结果表明，启动 300ms 时，词汇位置编码仍然具有灵活性，但首词素位置概率信息没有得到加工，因此不会影响词汇位置编码的灵活程度。

表 7 启动 300ms 时不同实验条件下的正确率和平均反应时（ms）

	正确率		反应时(ms)	
启动类型	高首词素位置概率	低首词素位置概率	高首词素位置概率	低首词素位置概率
原词启动	0.95(0.04)	0.93(0.06)	522(93)	525(89)
转换启动	0.95(0.04)	0.93(0.05)	672(107)	680(111)
替换启动	0.92(0.09)	0.91(0.08)	758(103)	755(102)

表 8 启动 300ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	*b*	*SE*	*t/z*	95%CI
正确率	截距	3.15	0.13	24.52***	[2.67, 3.04]
	首词素位置概率	-0.20	0.14	-1.44	[-0.45, 0.01]
	启动类型（UN vs TC）	**-0.48**	**0.17**	**-2.88****	**[-0.88, -0.32]**
	首词素位置概率×启动类型（UN vs TC）	0.23	0.33	0.68	[-0.65, 0.46]
反应时	截距	6.42	0.02	284.05***	[6.38, 6.47]
	首词素位置概率	0.01	0.01	0.80	[-0.01, 0.02]
	启动类型（UN vs TC）	**0.13**	**0.01**	**11.06*****	**[0.10, 0.15]**
	首词素位置概率×启动类型（UN vs TC）	-0.01	0.02	-0.71	[-0.05, 0.02]

2.4 讨论

实验 1 采用掩蔽启动词汇判断任务，通过操纵首词素位置概率高低和启动类型，在三个启动时间条件下考察首词素位置概率高低如何调节词汇位置编码的灵活程度。本实验有如下三方面发现：

第一，不同的启动时长影响中文词汇识别中的首词素位置概率效应，当启动时长为 150ms 时具有显著的首词素位置概率效应，表明首词素位置概率信息的加工主要发生在词汇加工中期（150ms 左右）。这与彭聃龄等人（1999）的研究结果一致，表明首词素位置概率信息作为词素表征的一部分，其激活经历了从无到有，再逐渐减弱的过程，其激活程度在词

汇加工中期最强。

第二，三个子实验均发现了稳定的词素转置效应，这与徐迩嘉和隋雪（2018）的研究发现一致，表明在词汇识别的早、中、晚期，双字词的位置编码均具有灵活性。对比三个启动时间条件下的词素转置效应大小，发现随着启动时间的增长，词素转置效应量逐步增大（80ms，150ms，300ms 启动时间的转置效应量分别为 38ms，60ms，81ms）。这表明词汇位置编码的灵活程度在整个词汇识别过程中逐步增强。而徐迩嘉和隋雪（2018）采用相同的实验范式，将无意义符号（ⅢⅢ）作为无关启动刺激，发现词汇位置编码的灵活程度经历了“弱-强-弱”的变化（80ms，150ms，300ms 启动时间的转置效应量分别为 91ms，167ms，117ms）。从词素转置效应量来看，徐迩嘉和隋雪（2018）在三个启动时间中的词素转置效应量均大于本研究（约为本研究的 2 倍），这可能是由于他们在研究中采用无意义符号作为无关启动刺激，目标词和无关启动刺激相似程度较低；而本研究用双字假词（首、尾字字频、笔画数、汉字结构与目标词相匹配）作为无关启动刺激，与目标词较为相似，导致词素转置效应量较小。由于两项研究对启动时间的操纵仅限于几个时间点，无法从加工时程的连续性视角反映词汇位置编码的灵活程度，且两项研究对无关启动刺激的操纵不一样，目前不清楚词汇位置编码灵活程度的不同变化模式是否和无关启动刺激的不同操纵相关。后续研究有必要借助脑成像技术（如 ERPs，MEG，或 fMRI），在不同的无关启动刺激条件下对词汇位置编码的加工时程，甚至于词汇位置编码灵活程度的变化时程进行进一步研究。

第三，首词素位置概率高低对词素转置效应的调节发生在词汇加工中期（150ms 左右）。具体表现为，当首词素经常用在词首时，词素转置效应量大于首词素不经常用在词首时。由于词素转置效应大小反映了词汇位置编码的灵活程度，这表明当一个词汇的首词素经常用在词首时，其位置编码的灵活程度更高。基于 Liang 等人（2015, 2017）的观点，当词素在词内的实际位置与其常用位置一致时，处于该位置的词素激活程度更高。从启动时间为 150ms 的反应时结果来看，高、低首词素位置概率词的词素转置效应差异主要体现在转换启动条件上。首词素位置概率低的词汇识别反应时（$M = 669$ms）显著长于首词素位置概率高的词汇（$M = 644$ms），这是导致首词素位置概率较高时替换启动条件与转换启动条件反应时差异较大的原因，即首词素经常用在词首时，词素转置效应更大。进一步解释为，首词素位置概率高的词汇（如“跟班”，“跟”位于词首的概率高，位于词尾的概率低）发生转置后，形成转换启动词“班跟”。由于转置词“班跟”并非真词，心理词典中没有对应的词条，读者只能自下而上地分别加工两个词素（“班”和“跟”），激活其位置概率信息（如“跟”经常用在词首）。当目标词“跟班”呈现时，“跟”在目标词内的位置（词首）与转换启动词激活的

词素位置概率信息（常用在词首）相一致，因此目标词加工较快，转换启动效应较大，位置编码灵活程度较高。同理，首词素位置概率低的词（如“值班”，“值”位于词首的概率低，位于词尾的概率则相对较高）发生转置后，激活其位置概率信息（如“值”经常用在词尾）。当目标词“值班”呈现时，“值”在目标词内的位置（词首）与转换启动词激活的词素位置概率信息（常用在词尾）冲突，读者需要额外时间解决冲突，导致转换启动效应更小，位置编码灵活程度更低。以往研究发现，词汇位置的灵活编码可能受以下两个因素影响:（1）词边界，相比于词内转置条件（深简居出），跨词边界转置条件下（喜走欢动）的位置编码灵活程度更低（顾俊娟 等，2020）。（2）词语形态结构，相比于偏正复合词，并列复合词和单语素词的汉字位置编码更为灵活（苏省之 等，2024）。本研究则发现了第三个影响词汇位置编码灵活程度的因素，即首词素的位置概率。这提示我们有必要在中文阅读模型（如 CRM）中加入词汇位置灵活编码机制以及调节过程，以提升模型的解释力。

3 实验 2　尾词素位置概率对词汇位置编码灵活程度的影响及时间进程

3.1 实验目的

采用掩蔽启动词汇判断任务，操纵尾词素的位置概率和启动类型，考察尾词素位置概率如何影响词汇位置编码的灵活程度。

3.2 方法

3.2.1 被试

实验 2a（启动 80ms）、2b（启动 150ms）和 2c（启动 300ms）各包括 48 名被试。共 144 名被试，平均年龄为 20.38 岁（*SD* = 1.82）。被试选择标准同实验 1。

3.2.2 实验设计

依据启动时间（分别为 80ms，150ms 和 300ms），分为三个子实验。每个子实验均采用 2（尾词素位置概率：高、低）×3（启动类型：原词启动、转换启动、替换启动）被试内设计。

3.2.3 实验材料

词素位置概率高低标准同实验 1。从 SUBTLEX-CH 语料库（Cai & Brysbaert, 2010）中选取 135 对低频且不可逆的双字词（平均词频：1.01 次/百万；词频范围：0.03 次/百万~5.35 次/百万）作为目标词，两类目标词的操纵和匹配方式同实验 1。操纵尾词素位置概率，高概率条件下，尾词素用在双字词词尾的平均概率为 84.62%（70.17%~100%）；低概率条件下，

尾词素用在双字词词尾的平均概率为 22.14%（5.26%~29.89%）。保证每对词的首词素相同，且首词素的位置概率处于中等水平（40.93%~59.57%）。匹配词频、尾词素字频、笔画数、整词笔画数。配对样本 t 检验结果显示，两类目标词的词频、尾词素字频、笔画数、整词笔画数均无显著差异，$ts < 1$，$ps > 0.05$。目标词材料的基本特征统计见表 9。

每个目标词有三种启动条件：原词条件、转换条件、替换条件。转换条件和替换条件下启动词的两个汉字在结构上匹配。转换启动词的首字字频和笔画数与替换启动词的首字无显著差异，$ts < 1$，$ps > 0.05$。转换启动词的尾字字频和笔画数与替换启动词的尾字无显著差异，$ts < 1$，$ps > 0.05$。

表 9 目标词材料基本统计特征

属性	高尾位置概率目标词	低尾位置概率目标词	转换启动词	替换启动词
整词词频	1.01(1.18)	1.00(1.16)	—	—
整词笔画数	17.17(4.29)	16.92(4.24)	17.04(4.26)	16.99(4.20)
首字字频	486.48(1106.77)	486.48(1106.77)	825.26(3903.27)	822.81(2290.62)
首字笔画数	8.58(3.11)	8.58(3.11)	8.47(3.01)	8.48(2.97)
尾字字频	635.75(1467.70)	1014.77(5325.25)	486.48(1104.71)	446.96(652.66)
尾字笔画数	8.59(3.03)	8.34(2.99)	8.58(3.11)	8.51(3.10)

加入 270 个假词作为填充，假词选择标准及评定同实验 1。目标词的首词素字频和笔画数（字频：$M = 486.48$ 次/百万，$SD = 1104.71$ 次/百万；笔画数：$M = 8.58$，$SD = 3.11$）与填充词无显著差异（字频：$M = 486.89$ 次/百万，$SD = 1121.17$ 次/百万；笔画数：$M = 8.59$，$SD = 3.19$；$ts < 1$，$ps > 0.05$）。目标词的尾词素字频和笔画数（字频：$M = 825.26$ 次/百万，$SD = 3903.27$ 次/百万；笔画数：$M = 8.47$，$SD = 3.01$）与填充词无显著差异（字频：$M = 681.40$ 次/百万，$SD = 1315.07$ 次/百万；笔画数：$M = 8.57$，$SD = 3.08$；$ts < 1$，$ps > 0.05$）。填充词材料的基本特征统计见表 10。

表 10 填充词材料基本特征统计

属性	填充词	目标词	转换启动词	替换启动词
首字字频	486.89(1121.17)	486.48(1104.71)	681.40(1315.07)	669.15(1257.28)
首字笔画数	8.59(3.19)	8.58(3.11)	8.57(3.08)	8.51(3.01)
尾字字频	681.40(1315.07)	825.26(3903.27)	486.89(1121.17)	487.15(986.16)

尾字笔画数	8.57(3.08)	8.47(3.01)	8.59(3.19)	8.62(3.11)
总笔画数	17.16(4.42)	17.04(4.26)	17.37(4.41)	17.13(4.43)

为防止被试在不同启动条件下看到同一目标词，将实验材料进行拉丁方分组，同实验 1。

3.2.4 实验仪器和实验程序

实验仪器和程序同实验 1，实验需 30 分钟左右。

3.3 结果

数据剔除方法和模型分析方法同实验 1，3 个实验删除数据分别占总数据的 4.45%、3.26%、6.84%。

3.3.1 实验 2a（启动时间为 80ms）的正确率和反应时分析

各个实验条件下的正确率和反应时结果见表 11 所示，模型统计结果汇总见表 12 所示。

正确率分析。尾词素位置概率主效应显著（b = -0.37, SE = 0.12, z = -3.06, p = 0.002），高尾词素位置概率词的正确率显著高于低尾词素位置概率词；转换启动与替换启动的正确率差异不显著（b = -0.09, SE = 0.13, z = -0.68, p = 0.50）；交互作用不显著（b = -0.01, SE = 0.26, z = -0.02, p = 0.98）。

反应时分析。尾词素位置概率主效应显著（b = 0.03, SE = 0.01, t = 5.37, p ＜ 0.001），低尾词素位置概率词的反应时显著长于高尾词素位置概率词；替换启动的反应时显著长于转换启动（b = 0.06, SE = 0.01, t = 8.06, p ＜ 0.001）；交互作用不显著（b = 0.02, SE = 0.01, t = 1.66, p = 0.10）。

以上结果表明，启动 80ms 时，尾词素位置概率高低影响词汇识别，此时位置编码具有灵活性，但尾词素位置概率不影响词汇位置编码的灵活程度。

表 11 启动 80ms 时不同实验条件下的正确率和平均反应时（ms）

	正确率		反应时(ms)	
启动类型	高尾词素位置概率	低尾词素位置概率	高尾词素位置概率	低尾词素位置概率
原词启动	0.93(0.04)	0.92(0.05)	610(92)	638(93)
转换启动	0.94(0.04)	0.91(0.06)	698(90)	704(88)
替换启动	0.94(0.05)	0.90(0.06)	729(89)	757(94)

表 12 启动 80ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	b	SE	t/z	95%CI
正确率	截距	2.74	0.09	31.57***	[2.57, 2.91]
	尾词素位置概率	**-0.37**	**0.12**	**-3.06****	**[-0.61, -0.13]**
	启动类型（UN vs TC）	-0.09	0.13	-0.68	[-0.35, 0.17]
	尾词素位置概率×启动类型（UN vs TC）	-0.01	0.26	-0.02	[-0.52, 0.51]

	截距	6.50	0.02	349.43***	[6.46, 6.53]
	尾词素位置概率	**0.03**	**0.01**	**5.37*****	**[0.02, 0.04]**
反应时	**启动类型（UN vs TC）**	**0.06**	**0.01**	**8.06*****	**[0.05, 0.08]**
	尾词素位置概率×启动类型（UN vs TC）	0.02	0.01	1.66	[-0.004, 0.05]

3.3.2 实验 2b（启动时间为 150ms）的正确率和反应时分析

各实验条件下的正确率和反应时结果见表 13 所示，模型统计结果汇总见表 14 所示。

正确率分析。尾词素位置概率主效应显著（$b = -0.65$, $SE = 0.10$, $z = -6.28$, $p < 0.001$），高尾词素位置概率词的正确率显著高于低尾词素位置概率词；替换启动的正确率显著低于转换启动（$b = -0.54$, $SE = 0.11$, $z = -4.75$, $p < 0.001$）；交互作用不显著（$b = 0.06$, $SE = 0.23$, $z = 0.28$, $p = 0.78$）。

反应时分析。尾词素位置概率主效应显著（$b = 0.03$, $SE = 0.01$, $t = 4.31$, $p < 0.001$），高尾词素位置概率词的反应时显著短于低尾词素位置概率词；替换启动的反应时显著长于转换启动（$b = 0.11$, $SE = 0.01$, $t = 14.22$, $p < 0.001$）；交互作用不显著（$b = -0.02$, $SE = 0.01$, $t = -1.58$, $p = 0.12$）。

上述结果表明，启动 150ms 时，尾词素位置概率高低影响词汇识别的正确率和速度。

	正确率		反应时(ms)	
启动类型	高尾词素位置概率	低尾词素位置概率	高尾词素位置概率	低尾词素位置概率
原词启动	0.95(0.04)	0.91(0.07)	523(78)	544(95)
转换启动	0.94(0.06)	0.91(0.08)	633(97)	655(92)
替换启动	0.91(0.09)	0.86(0.09)	705(88)	712(91)

表 13 启动 150ms 时不同实验条件下的正确率和平均反应时（ms）

表 14 启动 150ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	*b*	*SE*	*t/z*	95%CI
	截距	2.67	0.10	26.17***	[2.47, 2.87]
	尾词素位置概率	**-0.58**	**0.10**	**-6.09*****	**[-0.77, -0.40]**
正确率	**启动类型（UN vs TC）**	**-0.54**	**0.11**	**-4.74*****	**[-0.77, -0.32]**
	尾词素位置概率×启动类型（UN vs TC）	0.06	0.23	0.25	[-0.39, 0.50]
	截距	6.40	0.02	322.38***	[6.36, 6.44]
	尾词素位置概率	**0.03**	**0.01**	**4.31*****	**[0.02, 0.04]**
反应时	**启动类型（UN vs TC）**	**0.11**	**0.01**	**14.22*****	**[0.09, 0.12]**
	尾词素位置概率×启动类型（UN vs TC）	-0.02	0.01	-1.58	[-0.05, 0.01]

3.3.3 实验 2c（启动时间为 300ms）的正确率和反应时分析

各实验条件下的正确率和反应时结果见表 15 所示，模型统计结果汇总见表 16 所示。

正确率分析。尾词素位置概率主效应显著（b = -0.55, SE = 0.13, z = -4.31, $p < 0.001$），高尾词素位置概率词的正确率显著高于低尾词素位置概率词；替换启动的正确率显著低于转换启动（b = -0.42, SE = 0.14, z = -3.00, p = 0.003）；交互作用不显著（b = 0.11, SE = 0.28, z = 0.38, p = 0.71）。

反应时分析。尾词素位置概率主效应显著（b = 0.04, SE = 0.01, t = 4.57, $p < 0.001$），高尾词素位置概率词的反应时显著短于低尾词素位置概率词；替换启动的反应时显著长于转换启动（b = 0.12, SE = 0.01, t = 10.42, $p < 0.001$）；交互作用不显著（b = -0.01, SE = 0.02, t = -0.34, p = 0.74）。

上述结果表明，启动 300ms 时，尾词素位置概率影响词汇识别。

表 15 启动 300ms 时不同实验条件下的正确率和平均反应时（ms）

	正确率		反应时(ms)	
启动类型	高尾词素位置概率	低尾词素位置概率	高尾词素位置概率	低尾词素位置概率
原词启动	0.95(0.05)	0.92(0.06)	511(105)	533(111)
转换启动	0.94(0.06)	0.91(0.06)	668(116)	690(118)
替换启动	0.92(0.08)	0.88(0.09)	740(102)	765(102)

表 16 启动 300ms 时不同实验条件下正确率和反应时的线性混合模型统计结果

	固定效应	b	SE	t/z	95%CI
正确率	截距	2.89	0.12	23.43***	[2.64, 3.13]
	尾词素位置概率	**-0.55**	**0.13**	**-4.31*** **	**[-0.80, -0.30]**
	启动类型（UN vs TC）	**-0.42**	**0.14**	**-3.00** **	**[-0.69, -0.15]**
	尾词素位置概率×启动类型（UN vs TC）	0.11	0.28	0.38	[-0.44, 0.65]
反应时	截距	6.42	0.02	277.35***	[6.37, 6.46]
	尾词素位置概率	**0.04**	**0.01**	**4.57*** **	**[0.02, 0.05]**
	启动类型（UN vs TC）	**0.12**	**0.01**	**10.42*** **	**[0.10, 0.14]**
	尾词素位置概率×启动类型（UN vs TC）	-0.01	0.02	-0.34	[-0.04, 0.03]

3.4 讨论

实验 2 采用与实验 1 相同的实验范式，操纵尾词素位置概率高低和启动类型，在三个启

动时间条件下考察尾词素位置概率高低如何调节词汇位置编码的灵活程度。有如下三方面发现：

首先，尾词素位置概率的主效应在 3 种启动时间条件下均显著，表明该效应较为稳定，不受启动时间影响。这与 Liang 等人（2023）的研究结果相似，她们采用眼动追踪技术在句子阅读任务中考察首、尾词素位置概率信息的作用。在各项眼动指标上，均未发现首词素概率信息的主效应，表明成人读者和儿童读者均不能利用首词素位置概率信息进行词切分和词识别；而在反映词汇加工不同阶段的注视时间指标中（如首次注视时间、凝视时间和总注视时间），均发现尾词素位置概率信息的主效应显著，表现为当尾词素经常用在词尾时，目标词的注视时间显著变短。由于首次注视时间和凝视时间反映早期词汇通达，总注视时间反映后期词汇语义整合，不难看出，尾词素位置概率信息的作用时程更长、更稳定。

其次，实验 2 的三个子实验均发现词素转置效应，重复了实验 1 和徐迩嘉和隋雪（2018）的研究发现，再次表明词汇位置编码在整个词汇识别过程中均具有灵活性。再次对比三个启动时间条件下的词素转置效应大小，发现词素转置效应大小随启动时间的增长逐步增大（80ms，150ms，300ms 启动时间的转置效应量分别为 42ms，65ms，74ms），表明词汇位置编码的灵活程度在整个词汇识别过程中逐步增强。

最后，尾词素位置概率高低未调节词素转置效应大小，即尾词素位置概率高低不影响词汇位置编码的灵活程度。结合实验 1 以及实验 2 的第二方面研究发现（尾词素位置概率信息在中文词汇识别和句子阅读中的重要作用），似乎可以推断，首词素和尾词素在中文词汇识别中的作用方式可能不同：首词素位置概率信息调节词汇位置编码的灵活程度，但对词汇识别的影响程度较小（仅发生在影响词汇加工中期）；相比之下，尾词素位置概率信息虽然不会调节词汇位置编码的灵活程度，但在词汇识别中的作用较大。由于词素位置加工属于词汇识别的一部分，词汇识别不仅包括位置信息编码，还包括身份信息编码。由此推断，首词素的位置概率信息可能与词汇识别中的位置信息编码更为相关；尾词素的位置概率信息与词汇识别中的身份信息编码（如身份信息的激活强度）更为相关。

4 总讨论

本研究通过两个平行实验分别操纵首、尾词素位置概率高低和启动类型，考察了在不同启动时间下，首、尾词素位置概率如何影响双字词的位置编码灵活程度。综合两个实验，有三方面发现：（1）词汇位置编码在整个词汇识别过程中均具有一定程度的灵活性。（2）首、尾词素位置概率信息均作用于词汇识别，与首词素相比，尾词素的位置概率信息作用时程更

长。（3）词素位置概率信息对词汇位置编码灵活程度的调节表现出首、尾词素的差异，首词素位置概率调节词汇位置编码灵活程度，而尾词素位置概率则不影响词汇位置编码灵活程度。结合当前主流的中文阅读眼动控制模型，对上述发现进行讨论。

4.1 首、尾词素位置概率信息在词汇识别中的加工时程差异

综合实验 1 和实验 2 的研究结果，首词素位置概率效应受启动时长的影响，在启动时间为 150ms 时，该效应显著；而尾词素位置概率效应不受启动时长的影响，在三个启动时长中，该效应均显著。由此推断，与首词素相比，尾词素位置概率在中文词汇识别中的作用时程更长。需要说明的是，在分析首、尾词素位置概率主效应时，是将相同条件、转换条件和替换条件合并分析。相同启动条件类似于传统的词汇判断任务，在该条件下比较首、尾词素位置概率高低的加工时程，可发现，词汇判断时间不受首词素位置概率高低所影响（启动时间为 80ms、150ms、300ms 时，高、低首词素位置概率的平均反应时差异均为 3ms）；而当一个词汇的尾词素经常用在词尾时，该词汇的判断时间显著短于尾词素不经常用在词尾时（启动时间为 80ms、150ms、300ms 时，高、低首词素位置概率的平均反应时差异分别为 28ms、21ms、22ms）。上述结果表明，尾词素，而不是首词素的位置概率信息作用于中文词汇识别。这与 Liang 等人（2023）在中文阅读中的研究结果相同，即只有尾词素的位置概率信息作用于中文阅读的词切分和词识别。Liang 等人（2023）讨论到，这可能是由于当读者开始阅读一句话时，句首第一个词的左边界是确定的，读者只要依据尾词素的位置概率信息判断第一个词的结尾在哪里，就能完成该词汇的切分与识别。由于中文阅读没有空格，相邻两个词共享一个词边界，即上一词的右边界是下一词的左边界，因此，上一个词的词尾确定后，便可以同时确定下一个词的词首边界。基于认知加工的经济性原则，读者无需再利用首词素的位置概率信息切分词首，只需利用尾词素的位置概率信息判断词尾的位置即可完成词汇识别。本研究发现，在中文词汇识别中，首词素位置概率信息的加工时程较短（或者说是“无”），尾词素位置概率信息加工时程较长（从 80ms 持续至 300ms），也可能是由于在词汇识别中，首字位于最左侧，其身份和位置是相对固定的，读者只需要借助于词尾的位置概率信息判断首字和尾字组合起来是否可以通达，便能做出词汇判断。如果该解释成立，我们有理由推断，首、尾词素的位置概率信息在中文阅读和词汇识别情境中的作用方式似乎是相同的。

不过，曹海波等人（2023）在一个实验中同时操纵首、尾词素的位置概率信息，却发现在低频词识别中，首词素，而不是尾词素的位置概率信息作用于词汇识别，这与本研究和 Liang 等人（2023）的结果相反。由于曹海波等人（2023）和本研究最大的不同体现在核心

变量的操纵方式上：本研究在操纵首词素位置概率高低时保持尾词素一致，在操纵尾词素位置概率高低时保持首词素一致；而曹海波等人（2023）在同一实验中同时操纵首、尾词素位置概率高低。由于首词素本身在词汇识别中的作用相对较大，也具有注意优势，因此，曹海波等人的操纵实则是在保证首词素作用的前提下，对首、尾词素位置概率信息的作用进行考察。本研究则是在保证首词素作用相同的条件下（注意优势相同），考察尾词素位置概率信息的作用，这种操纵方式更容易凸显尾词素在词汇识别中的作用。后续研究有必要在同一研究中，比较两种不同操纵方式下首、尾词素位置概率信息在中文词汇识别和阅读中的作用。

4.2 首、尾词素位置概率信息对词汇位置编码灵活程度的调节差异

实验 1 和实验 2 均发现，在词汇识别的不同阶段（早、中、晚期），词汇位置编码均具有灵活性。基于中文阅读模型（简称 CRM，见 Li & Pollatsek, 2020）的基本观点，对词汇位置编码的灵活机制进行解释。需要说明的是，由于在本研究所设置的最短启动时间（80ms）内，被试能够激活知觉广度范围内汉字的视觉信息，因此本研究在启动时间长于 80ms 后所发现的词素转置效应可能更多地对应于中文阅读模型的汉字和词汇加工层级：首先，在汉字加工层级，一个汉字被激活的同时，也会激活由它组成的所有词汇。在已激活的所有词汇中，该汉字可能位于词汇的不同位置（词首或词尾），此时，词汇的位置编码较为灵活；在词汇加工层级，已有的词汇表征和汉字加工层级的前馈激活交互作用于词汇识别。由于本研究的目标词均为低频词，主要依赖词素通达（白学军 等, 2015），汉字加工层级所表现出的词汇位置灵活编码可能延续到词汇加工层级，表现为在词汇通达晚期位置编码仍具有一定程度的灵活性。然而，本研究发现的词汇位置编码灵活程度的变化模式（早期-中期-晚期，灵活程度逐步增强）与徐迩嘉和隋雪（2018）研究（早期-中期-晚期，灵活程度先增强后减弱）不同，具体表现为：在词汇加工早期到中期，两项研究均发现词汇位置编码的灵活程度逐步增强。从词汇加工早期到中期，汉字的身份信息和位置信息的作用不断加强，由汉字组成的所有词汇被不同程度激活，在所有激活的词汇中，汉字位于词内的不同位置，因此词汇位置信息编码灵活程度逐步增强，这符合中文阅读模型的基本观点。在词汇加工中期到晚期，本研究发现词汇位置编码灵活程度逐步增强，而徐迩嘉和隋雪（2018）发现词汇位置编码灵活程度逐步减弱。也就是说，目前关于从词汇加工中期到晚期位置编码灵活程度如何变化还未得到统一定论。由于词汇加工的中、晚期更多地涉及词汇层和汉字层的交互作用，后续研究有必要进一步考察自上而下的词汇表征和自下而上的词素表征如何交互作用于词汇位置的灵活编码。

此外，本研究还发现，在启动时间为 150ms 时，双字词首词素的位置概率信息会调节

词汇位置编码的灵活程度，当词首的汉字经常用在词首时，词汇位置编码更灵活。但是，尾词素的位置概率信息没有调节词汇位置编码的灵活程度。这是本研究的主要发现。基于词汇位置编码的灵活程度与词汇识别的难易程度有关，位置编码越灵活，词汇识别越容易，越倾向于整词通达（Gu & Li, 2015），本研究发现表明当首词素的位置信息得到加工时，其位置概率高低通过影响词汇位置加工的灵活程度调节词汇识别的难易程度；但是当尾词素的位置信息得到加工时，其位置概率高低直接影响词汇识别，而不通过影响词汇位置加工的灵活程度调节词汇识别的难易。进一步表明首词素的位置概率信息和词汇位置编码的灵活程度有关，但尾词素的位置概率信息则和词汇位置编码的灵活程度无关。这可能是由于首词素在双字词识别中具有注意优势和加工优势（Ma & Li, 2015; Taft, 2004; Yan et al., 2006），其位置信息更重要。特别是一个词汇的首词素经常用在词首时，首词素的注意和加工优势更容易显现，词汇识别也就更容易，由此表现出词汇位置编码更加灵活的特点。相比之下，虽然尾词素的位置信息在双字词识别中不具有注意和加工优势，但是中文读者在阅读中主要利用尾词素所提供的信息进行词切分和词识别（Liang et al., 2023），这在一定程度上弱化其位置信息，而凸显其身份信息在词汇识别中的作用。

综上所述，首、尾词素位置概率信息在词汇识别中的作用以及对词汇位置编码灵活程度的调节有差异：首词素位置概率信息的加工时程较短，但与词汇位置编码的灵活程度有关；尾词素位置概率信息的加工时程较长，但与词汇位置编码的灵活程度无关，可能和词汇身份信息加工相关。这提示在中文阅读模型的词汇加工模块，有必要纳入首、尾词素的位置概率信息，以及二者在词汇识别身份信息加工和位置信息加工中的不同作用，进一步增强模型的解释力。

5 结论

本研究得出如下结论：（1）词汇位置编码的灵活性体现在整个词汇识别过程中。（2）首、尾词素位置概率信息均作用于词汇识别，相比于首词素，尾词素的位置概率信息作用时程更长。（3）首词素而不是尾词素的位置概率信息调节词汇位置编码灵活程度，当首词素经常用在词首时，词汇位置编码更灵活。支持首、尾词素位置概率信息在词汇识别中的作用不同的观点。

参考文献

Baciero, A., Gomez, P., Duñabeitia, J. A., & Perea, M. (2022). Raeding with the fingres: Towards a universal model of letter position coding. *Psychonomic Bulletin & Review, 29*(6), 2275−2283.

Bai, X. J., Li, X., & Yan, G. L. (2015). Eye movement controlin Chinese reading: A summary over the past 20 years ofresearch. *Psychological Development and Education, 31*(1),85–91.

[白学军, 李馨, 闫国利. (2015). 汉语阅读眼动控制: 20 年研究的总结. *心理发展与教育, 31*(1), 85–91.]

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2023). *lme4: Linear mixed−effects models using 'Eigen' and S4*. Retrieved July 4, 2023, from https://cran.r-project.org/web/packages/lme4/index.html

Blythe, H. I., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2014). Reading transposed text: Effects of transposed letter distance and consonant-vowel status on eye movements. *Attention, Perception, & Psychophysics, 76*(8), 2424−2440.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX−CH: Chinese word and character frequencies based on film subtitles. *PloS One, 5*(6), e10729.

Cao, H. B., Lan, Z. B., Gao, F., Yu, H. T., Li, P., & Wang, J. X. (2023). The role of character positional frequency on word recognition during Chinese reading: Lexical decision and eye movements studies. *Acta Psychologica Sinica, 55*(2), 159−176.

[曹海波, 兰泽波, 高峰, 于海涛, 李鹏, 王敬欣. (2023). 词素位置概率在中文阅读中的作用:词汇判断和眼动研究. *心理学报, 55*(2), 159−176.]

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review, 117*, 713−758.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review, 115*(3), 577−600.

Gu, J. J., Li, X. S., & Liversedge, S. P. (2015). Character order processing in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 41*(1), 127−137.

Gu, J. J., Gao, Z. H., & Qu, Q. Q. (2020). The word boundary effect of Chinese character position processing. *Studies of Psychology and Behavior, 18*(2), 193−199.

[顾俊娟, 高志华, 屈青青. (2020). 汉字位置加工的词边界效应.*心理与行为研究, 18*(2), 193−199.]

Hansell, M. (2008). Teaching and learning Chinese as a foreign language: a pedagogical grammar. *Modern Language Journal, 92*(2), 331−332.

Hua, H. M., Gu, J. J., Lin, N., & Li, X. S. (2017). Letter/character position encoding in visual word recognition. *Advances in Psychological Science, 25*(7), 1132−1138.

[滑慧敏, 顾俊娟, 林楠, 李兴珊. (2017). 视觉词汇识别中的字符位置编码. *心理科学进展, 25*(7), 1132−1138.]

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye

movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 209−229.

Li, X. S., & Pollatsek, A. (2020). An integrated model of word processing and eye−movement control during Chinese reading. *Psychological Review, 127*(6), 1139−1162.

Lian, K. Y., Ma, J., Wei, L., Zhang, S. W., & Bai, X. J. (2021). The role of character positional frequency on college and primary student in oral reading. *Studies of Psychology and Behavior, 19*(2), 179−185.

[连坤予, 马杰, 魏玲, 张书袆, 白学军. (2021). 汉语朗读中词素位置概率线索作用的发展研究. *心理与行为研究, 19*(2), 179−185.]

Liang, F. F., Blythe, H. I., Bai, X. J., Yan, G. L., Li, X., Zang, C. L., & Liversedge, S. P. (2017). The role of character positional frequency on Chinese word learning during natural reading. *PloS One, 12*(11), e0187656.

Liang, F. F., Blythe, H. I., Zang, C. L., Bai, X. J., Yan, G. L., & Liversedge, S. P. (2015). Positional character frequency and word spacing facilitate the acquisition of novel words during Chinese children's reading. *Journal of Cognitive Psychology, 27*(5), 594−608.

Liang, F. F., Feng, L. L., Liu, Y., Li, X., & Bai, X. J. (2024). Different roles of initial and final character positional probabilities on incidental word learning during Chinese reading. *Acta Psychologica Sinica, 56*(3), 281−294.

[梁菲菲, 冯琳琳, 刘瑛, 李馨, 白学军. (2024). 阅读伴随词汇学习的词切分: 首、尾词素位置概率的不同作用. *心理学报, 56*(3), 281−294.]

Liang, F. F., Gao, Q., Li, X., Wang, Y. S., Bai, X. J., & Liversedge, S. P. (2023). The importance of the positional probability of word final (but not word initial) characters for word segmentation and identification in children and adults' natural Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49*(1), 98−115.

Liang, F. F., Liu, Y., Feng, L. L., He, F., Wang, Z., & Bai, X. J. (2024). Visual complexity effect in Chinese incidental word learning: Evidence from number of strokes and word length. *Acta Psychologica Sinica, 56*(12), 1734−1750.

[梁菲菲, 刘瑛, 贺斐, 冯琳琳, 王峥, 白学军. (2024). 中文阅读伴随词汇学习中的视觉复杂性效应：基于笔画数和词长的证据.*心理学报,56*(12), 1734−1750.]

Ma, G. J., & Li, X. S. (2015). How character complexity modulates eye movement control in Chinese reading. *Reading and Writing, 28*(6), 747−761.

Peng, D. L., Ding, G. S., Wang, C. M., Taft, M., & Zhu, X. P. (1999). The process of Chinese reversible words-the role of morphemes in lexical access. *Acta Psychologica Sinica, 31*(1), 36−46.

[彭聃龄, 丁国盛, 王春茂, Taft, M., 朱晓平. (1999). 汉语逆序词的加工——词素在词加工中的作用. *心理学报, 31*(1), 36−46.]

Perea, M., & Fraga, I. (2006). Transposed-letter and laterality effects in lexical decision. *Brain and Language, 97*(1), 102−109.

Perea, M., Winskel, H., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: The case of Thai. *Experimental Psychology, 59*(2), 68−73.

R Development Core Team. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Song, X. N., Xu, X. C., Yang, X. L., Sun, G. L., & Cui, L. (2022). The Influence of Predictability, Word Frequency and StrokeNumber on Chinese Word Recognition:An Eye Movement Study. *Journal of Psychological Science, 45*(5), 1061−1068.

[宋悉妮, 徐晓晨, 杨秀莉, 孙桂苓, 崔磊. (2022). 预期性、词频和笔画数对中文词汇识别影响的眼动研究. *心理科学, 45*(5), 1061−1068.]

Su, X. Z., Li, X. X., Li, R. R., Zhao, C. Z., & Cui, L. (2024). Morphological structures of two-character words influence character position encoding. *Acta Psychologica Sinica, 56*(4), 383−393.

[苏省之,李骁轩,李蓉蓉, 赵长泽, 崔磊. (2024). 双字词的形态结构对汉字位置信息编码的影响. *心理学报, 56*(4), 383–393.]

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology, 57*(4), 745–765.

Xu, E. J., & Sui, X. (2018). Effects of predictability on the time course of identity information and location information in Chinese word recognition. *Acta Psychologica Sinica, 50*(6), 606–621.

[徐迩嘉, 隋雪. (2018). 身份信息与位置信息的加工进程及语境预测性的影响. *心理学报, 50*(6), 606–621.]

Yan, G. L., Tian, H. J., Bai, X. J., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97*(2), 259−268.

Zhang, Y. C., Jiang, Y., Dong, Q., & Wang, J. X. (2021). The identity and position for Chinese character, digit, and symbol strings-special mechanism of Chinese reading. *Psychological Exploration, 41*(6), 504−514.

[张妍萃, 江莹, 董青, 王敬欣. (2021). 汉字、数字和符号的身份信息和位置信息加工比较——中文阅读的特别机制. *心理学探新, 41*(6), 504−514.]

The Influence of Initial and Final Character Positional Probabilities on the Flexibility of Position Coding in Two-Character Word Identification

LI Xin[1], ZHANG Mei[1], GU Junjuan[2], WANG Yongsheng[2], LIANG Feifei[1]

([1] *Faculty of Psychology, Tianjin Normal University, Tianjin* 300387, *China*) ([2] *School of Psychology and Mental Health, North China University of Science and Technology, Tangshan* 063210, *China*)

Abstract

Letter/character position information plays an important role in visual word identification. In alphabetic and Chinese reading, a bulk of studies using the lexical decision paradigm have shown that a transposed letter/character (TL/TC) nonword prime, created by transposing two adjacent letters/characters of a word, activates the representation of its base word to a greater degree than a substituted letter/character (SL/SC) nonword prime, in which two corresponding letters/characters, or even one letter/character are replaced. This phenomenon is called the Transposed Letter / Character effect, indicating that the letter/character position encoding is not strict during word identification. In alphabetic languages (like English, German and Spanish), letter order encoding mechanism have been included in the models of word recognition in different ways. Chinese is a logographic language, with quite different characteristics from alphabetic languages, so the letter position encoding models developed from alphabetic languages might not be fully applied to Chinese. In the present study, we conducted two parallel experiments to examine whether and how character positional probabilities influenced the degree of flexibility of position encoding in two-character word identification. We utilized a masked priming paradigm lexical judgment task and manipulated the priming type (original word priming, transposed word priming, substitution priming) and priming times (80ms, 150ms, 300ms).

In Experiment 1, we examined how word’s initial character's positional probability influenced the degree of flexibility of position coding in two-character word identification. The initial character’s positional probability of two-character target words was manipulated as being either high or low, and the final character was kept identical across the two conditions. In Experiment 2, an analogous manipulation was made for the final character of the target word to check whether the final character's positional probability of two-character words influenced the degree of flexibility of position coding in two-character word identification.

We found that the positional probability of both the initial and final character of a word influences Chinese two-character word identification. To be somewhat different, the positional probability of final character was processed throughout the whole lexical identification (from 80ms to 300ms), whereas the positional probability of initial character was processed only in the middle stage of lexical identification (around 150ms). This finding suggests that the final character's positional probability is more important than that of initial character’s during Chinese two-character word identification. Furthermore, the positional probability of the initial character rather than the final character modulated the degree of flexibility of position coding in two-character word identification, such that the more frequently a character being used as word beginning, the more flexible of position coding.

Based on these findings, we argue that the role of positional probability of initial and final character might be somewhat different during Chinese word identification: the positional probability of initial character may be related to positional information processing, whilst the positional probability associated with the final character might be more related to identity information processing in word identification.

Keywords transposed-character effect, character positional probability, character position encoding, two-character word identification